



# Cloud Storage for Science

**Massimo Lamanna**

**On behalf of the CERN IT Storage group (\*)**

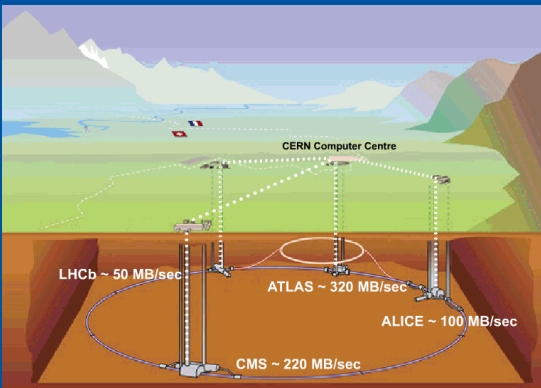
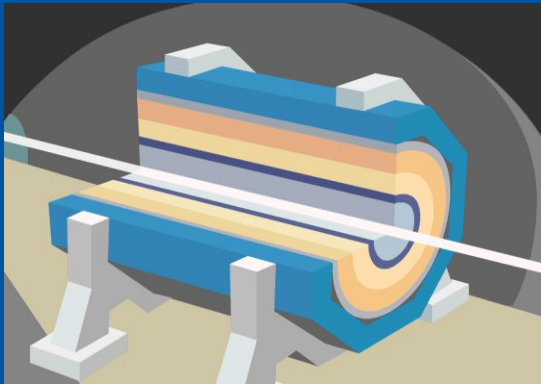
***CERN – Soleil visit (22-JUN-2016)***

# Outline

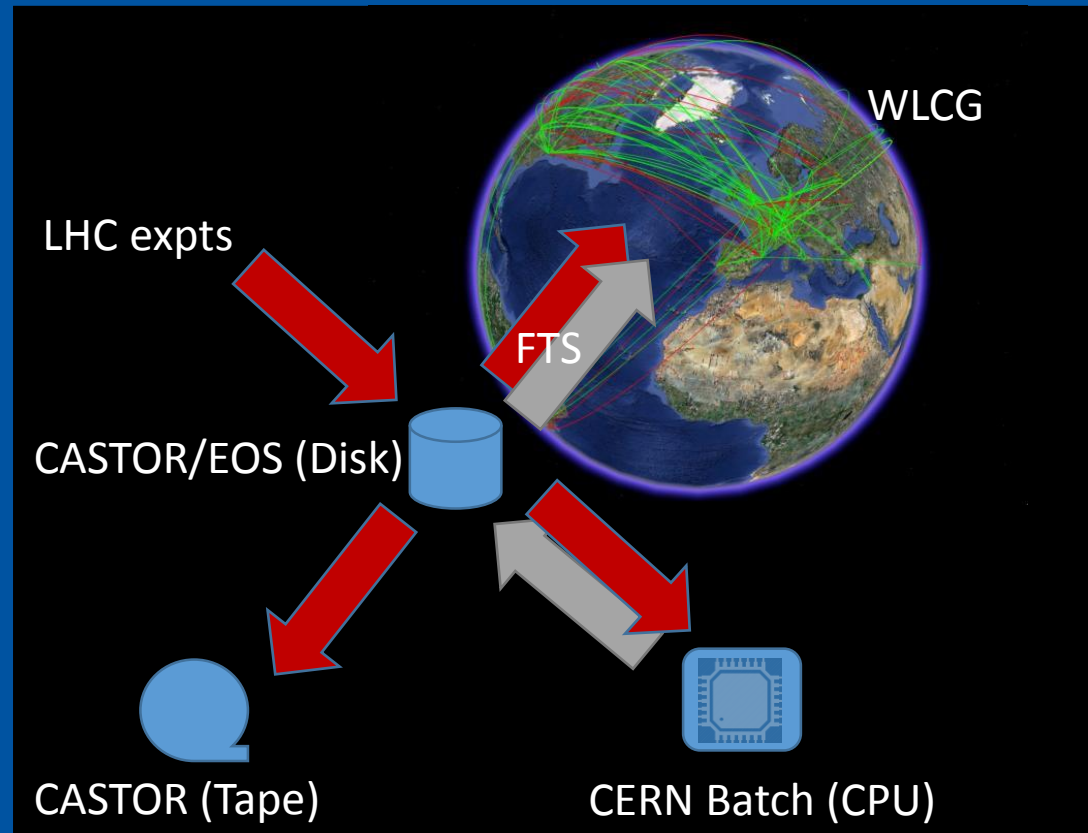
- Storage at CERN
  - Short recap
- EOS
  - Disk solution for LHC
- CERNBox
  - Cloud storage services
    - based on EOS and ownCloud

# Storage at CERN

# WLCG computing



- Reconstruction use case
  - Goals: data quality and immediate access for analysis
  - Organised activity dominated by heavy processing and replication
- Analysis use case
  - Goals: extract physics quantities (discovery)
  - Individual activities dominated by event selection and sharing



# IT ST data services

## AFS

- Home directories (user quota up to 100 GB)
- O(1) PB
- Backup

## CASTOR

- Tape archive with disk staging area
- O(100) PB archive

## EOS

- Disk farm
- O(100) PB disk space

## EOS/CERNBox

- User space (sync client; initial user quota 1 TB)
- AFS use cases for new/large projects
- Backup user space

## Ceph

- Virtualised infrastructure foundation
- CERN has a seat in the Ceph Advisory Board
- Block storage (OpenStack), Object store, ...

## FILER

- High-performance NFS servers

TSM, HADOOP, ...



YOU WANT YOUR COUSIN TO SEND YOU A FILE? EASY.  
HE CAN EMAIL IT TO— ... OH, IT'S 25 MB? HMM...  
DO EITHER OF YOU HAVE AN FTP SERVER? NO, RIGHT.  
IF YOU HAD WEB HOSTING, YOU COULD UPLOAD IT...  
HMM. WE COULD TRY ONE OF THOSE MEGASHARE/UPLOAD SITES,  
BUT THEY'RE FLAKY AND FULL OF DELAYS AND PORN POPUPS.  
HOW ABOUT AIM DIRECT CONNECT? ANYONE STILL USE THAT?  
OH, WAIT, DROPBOX! IT'S THIS RECENT STARTUP FROM A FEW  
YEARS BACK THAT SYNCs FOLDERS BETWEEN COMPUTERS.  
YOU JUST NEED TO MAKE AN ACCOUNT, INSTALL THE—

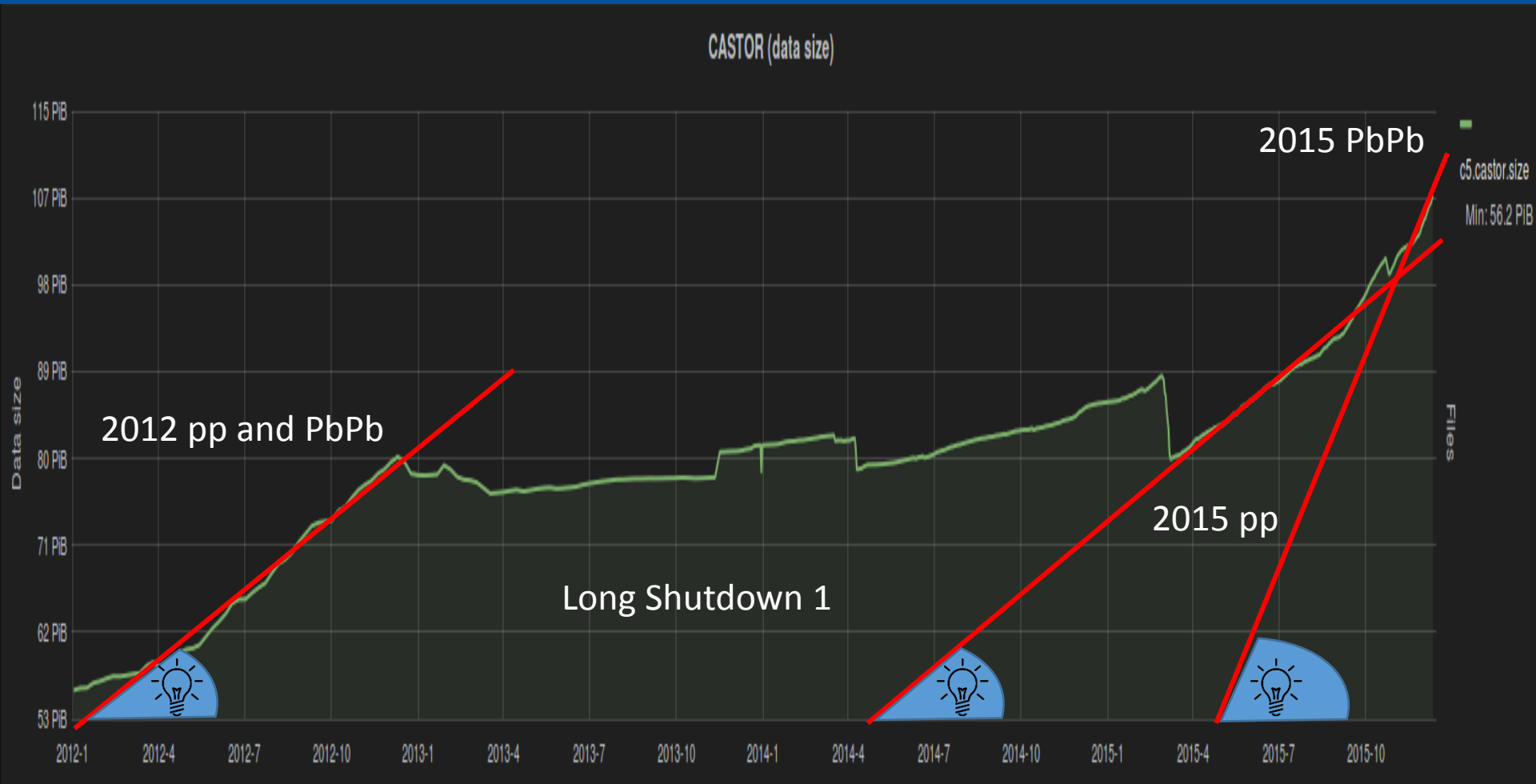


I LIKE HOW WE'VE HAD THE INTERNET FOR DECADES,  
YET "SENDING FILES" IS SOMETHING EARLY  
ADOPTERS ARE STILL FIGURING OUT HOW TO DO.

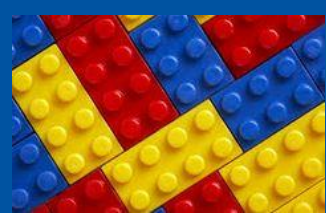
Randall Munroe (<http://xkcd.com>)



# (LHC) Data Taking 2012 → 2015

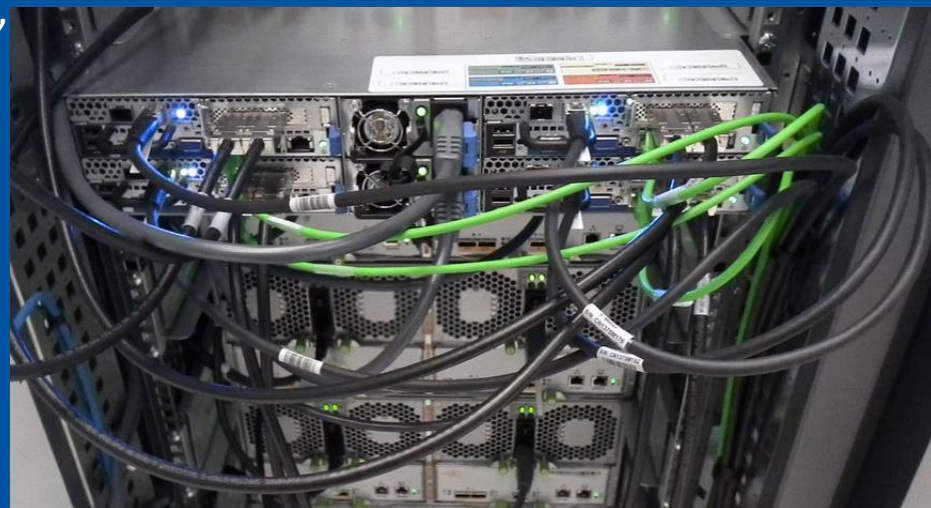


Last week record: 2 PB in a week



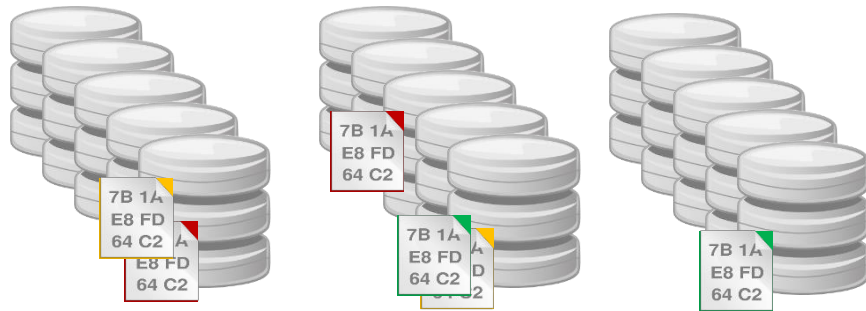
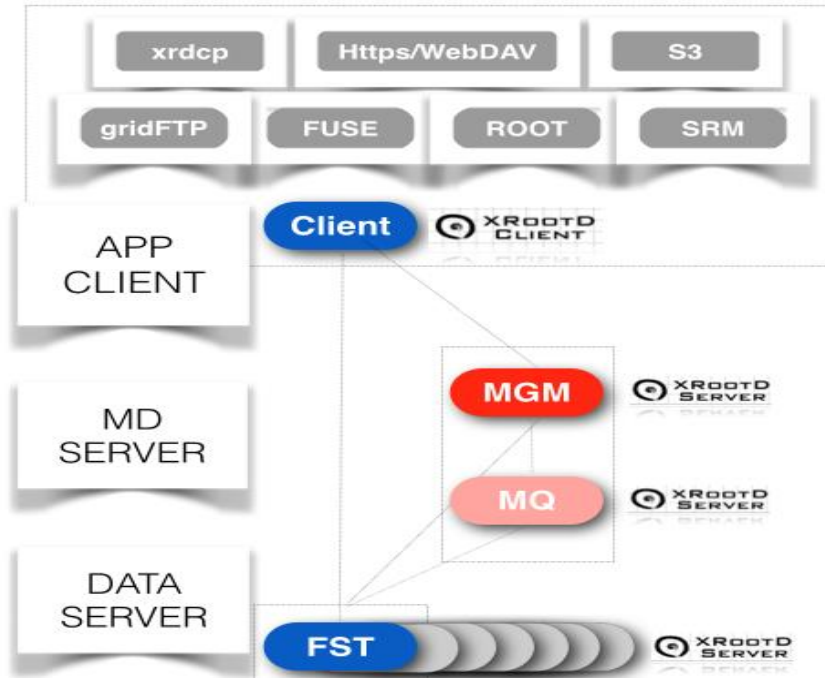
# Commodity / Uniformity

- Uniform HW across services
  - Same building blocks for AFS, CASTOR, Ceph and EOS
  - AFS:CASTOR:EOS 1:10:70 ratio
  - Heterogeneity essentially from technology evolution
    - Yearly cycle; new models with best “price per GB”
  - Recently took over NetApp NFS services
    - Now on the same hardware than EOS & co.
- Main building block
  - $\frac{1}{4}$  of quad system (1 node with dual CPU, 64 GB RAM, 10Gbit/s NIC)
    - OpenStack hypervisor use the same hardware
  - 2 24-disk trays per node
  - 24 6-TB disks per tray
    - 55 PB delivery in 2016Q2
  - → 200 TB building block





EOS



› EOS: Large disk farms for physics and beyond

- Towards 190 PB JBOD

› Developed in CERN/IT

› Original goal

- Large scale (PBs for 100s/1000s independent scientists) analysis of LHC data

› Strategic points

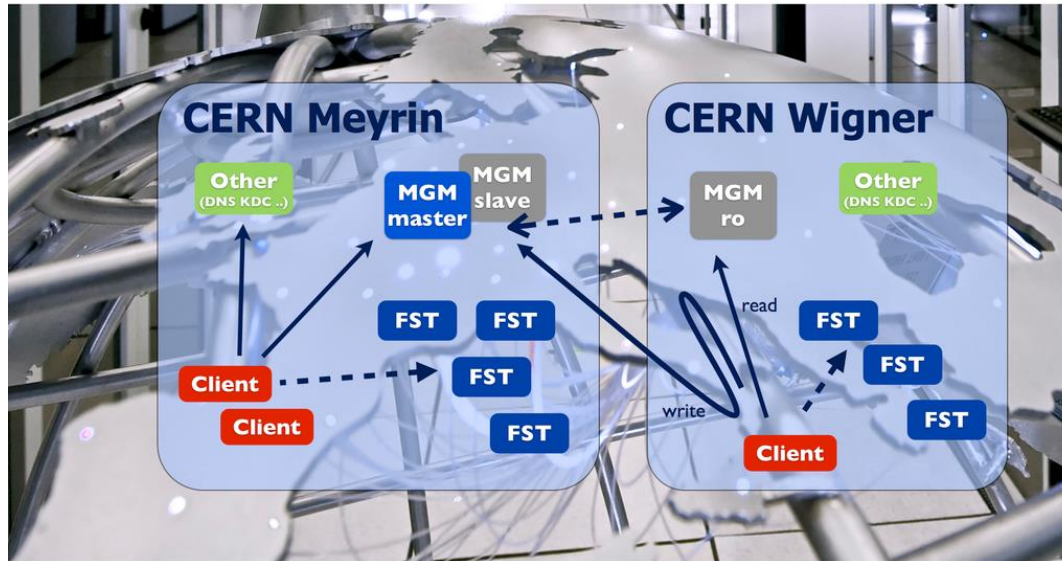
- Distill 20+ years of experience in HEP/data management
- Ultra-fast name space
- Optimised protocols
- Arbitrary level of data durability via cross-node file replication or RAIN using commodity hardware (complete control on the JBOD layer)

› Status

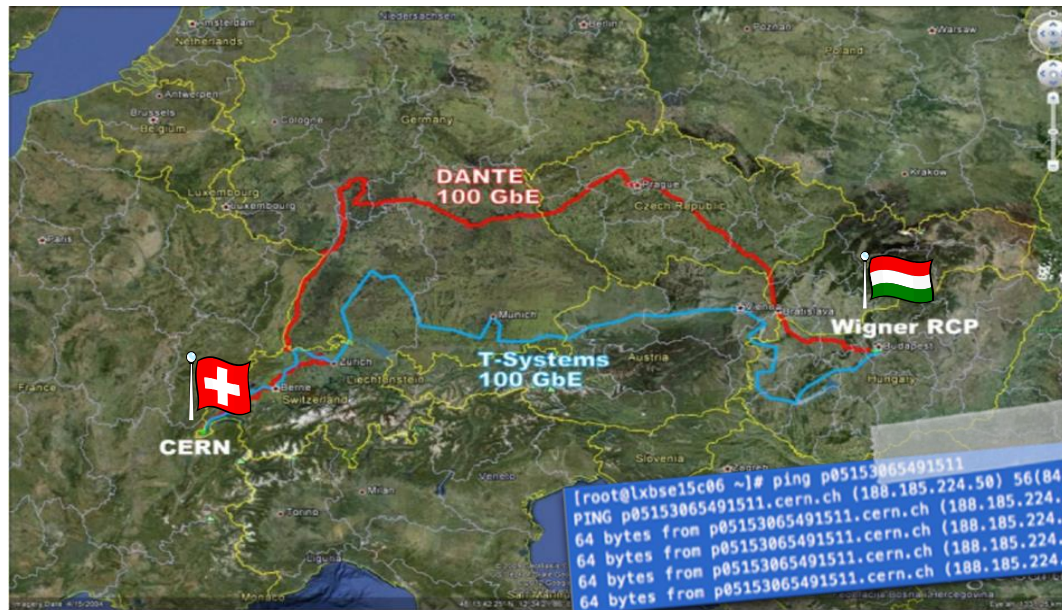
- Open to non-physics use cases
- Large number of protocols available
- “All” client platform
  - Red Hat Linux + OSX
  - (More Linux + Windows)
- Catalogue scale up:
  - $10^8$ - $10^9 \rightarrow 10^{10}$ - $10^{11}$



# Our “20-ms-large” computer centre



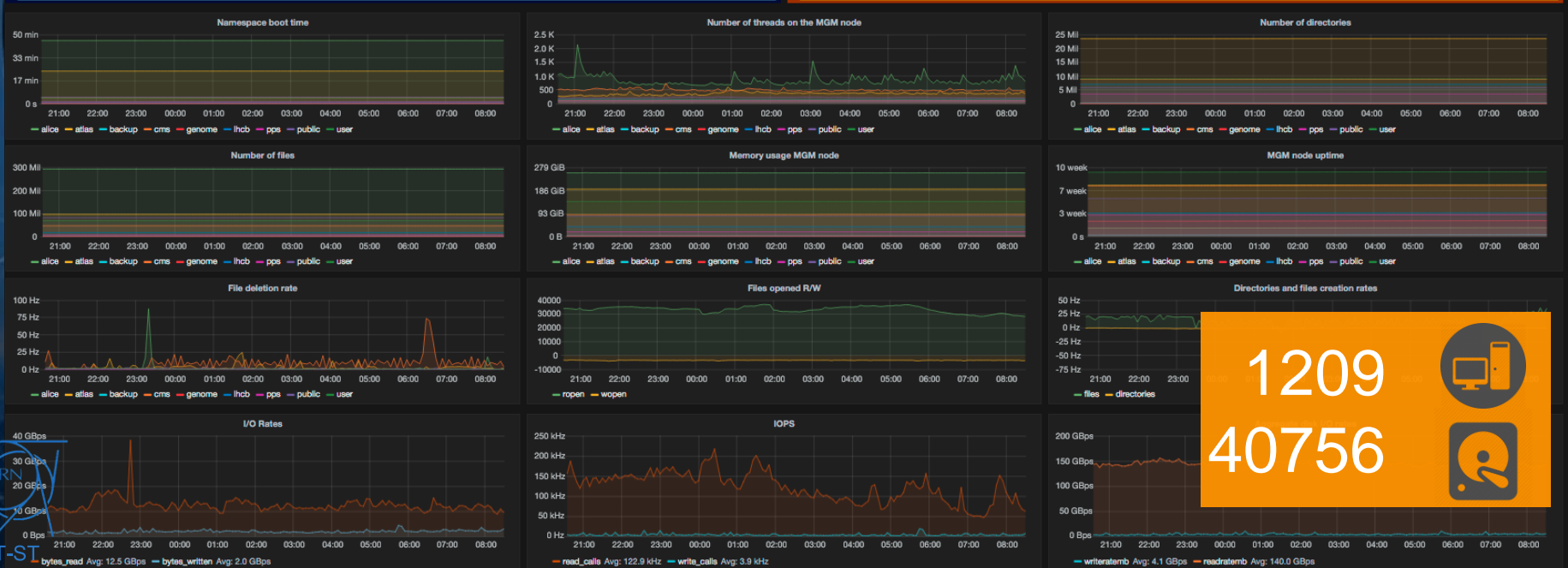
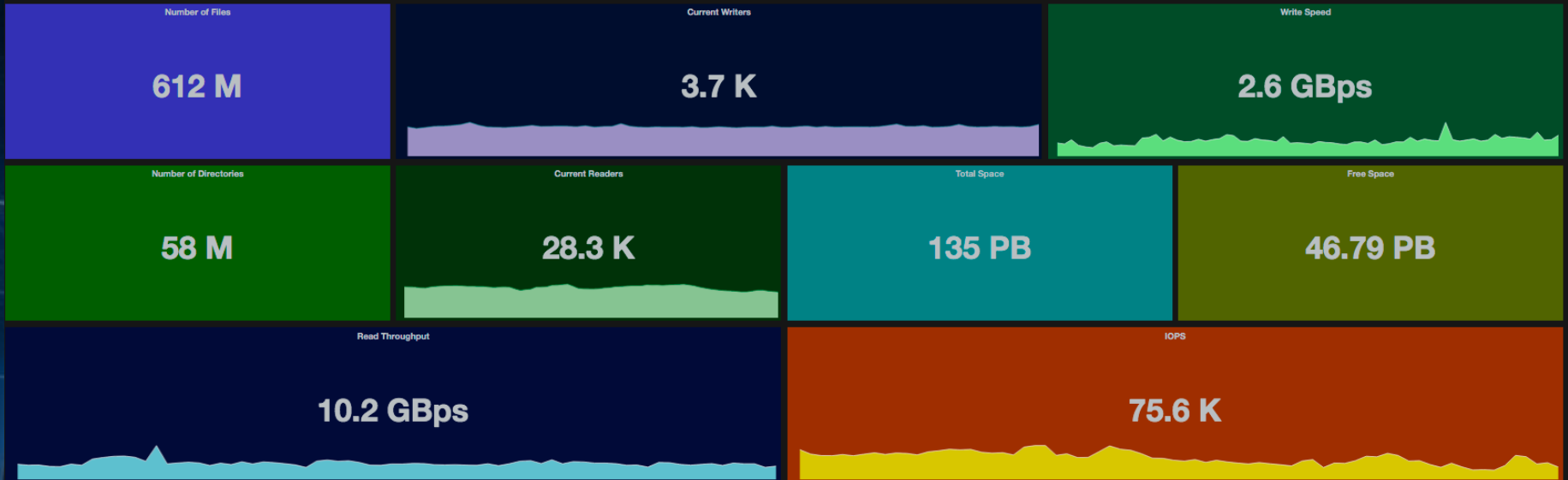
Autonomic,  
Locality,  
Disaster recovery/  
business continuity



MGM = NameSpace/Metadata FST = Disk servers



# EOS (April 2016)



# CERNBox

# Rationale behind cloud storage



- Existing professional usage of public cloud storage (Dropbox) at CERN
  - CERN (IT)
    - Harmonically extend our service portfolio
  - Today users expect:
    - Cross-platform, easy install and setup (BYOD)
    - Modern web interfaces
    - Mobile access
  - Selected ownCloud as starting point
    - Open software; data on premise
    - General public use and deployments in edu/research

• **Scale up in size:** interesting challenge (leverage on EOS)

• **Extend in the use-case phase space:** innovation



# What is CERNBox ?

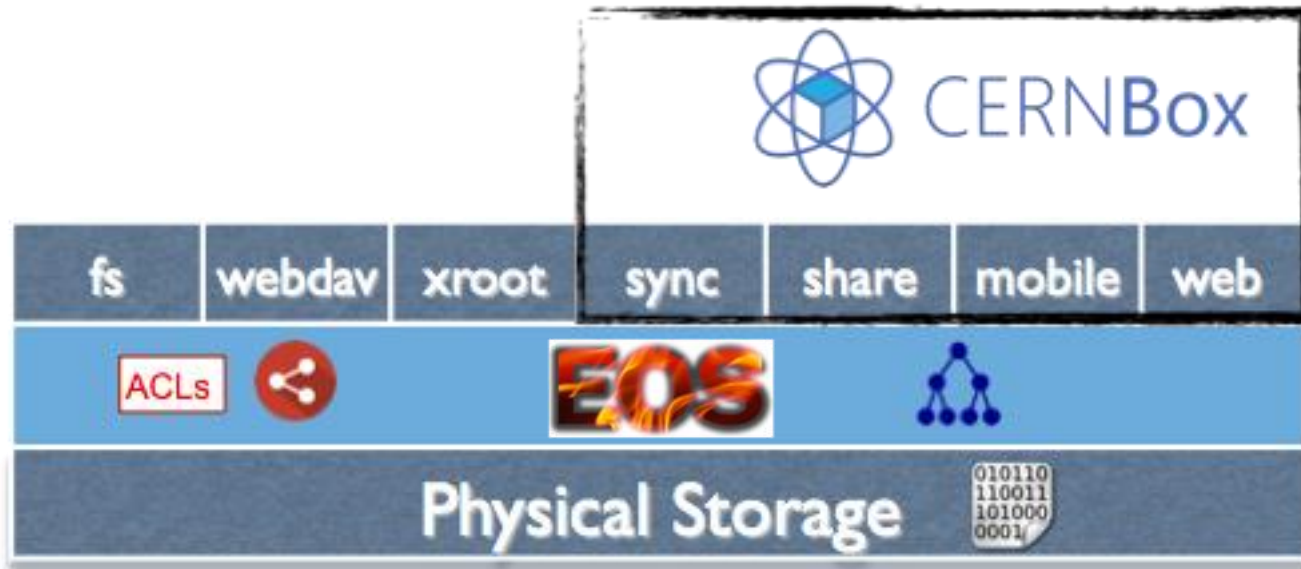


CERNBox provides a cloud synchronisation service

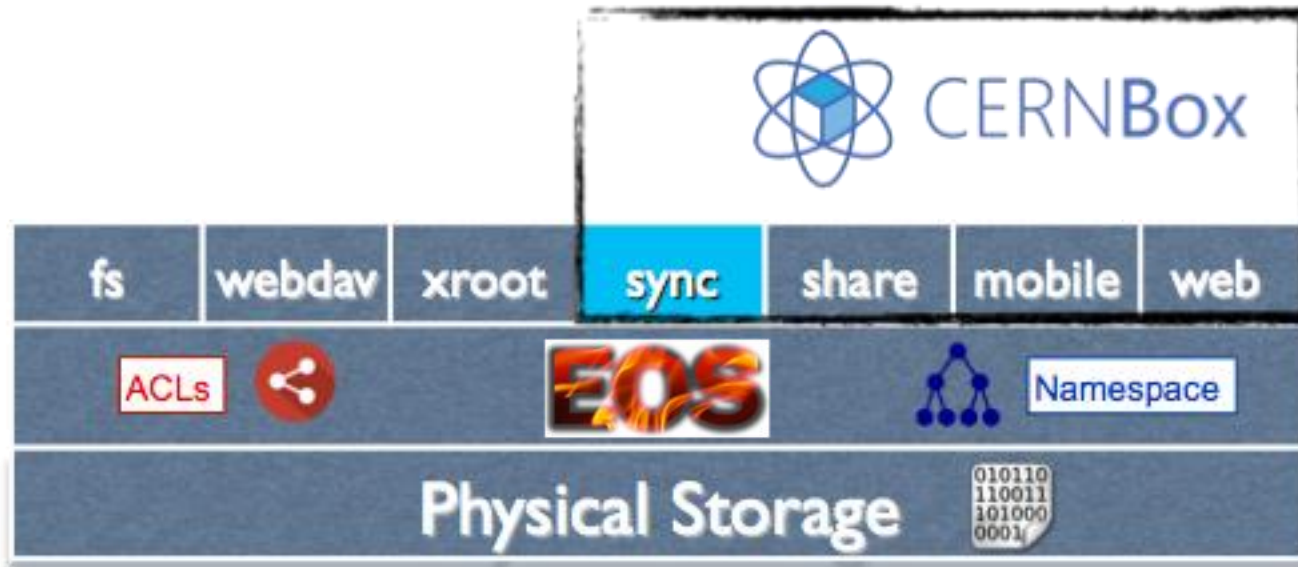
- Synchronise files (data at CERN) and offline data access
- Easy way to share with other users
- All major platforms supported
- Based on ownCloud integrated with EOS
  - Available for all CERN users (1TB/user initial quota)



Much more than a Dropbox™ replacement!

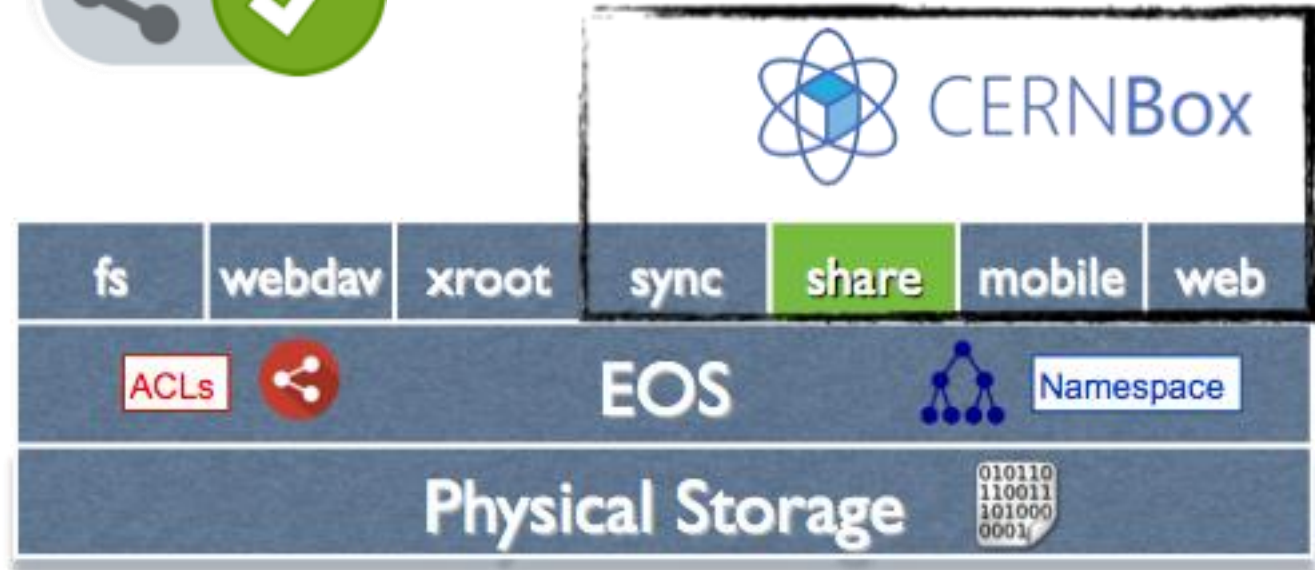
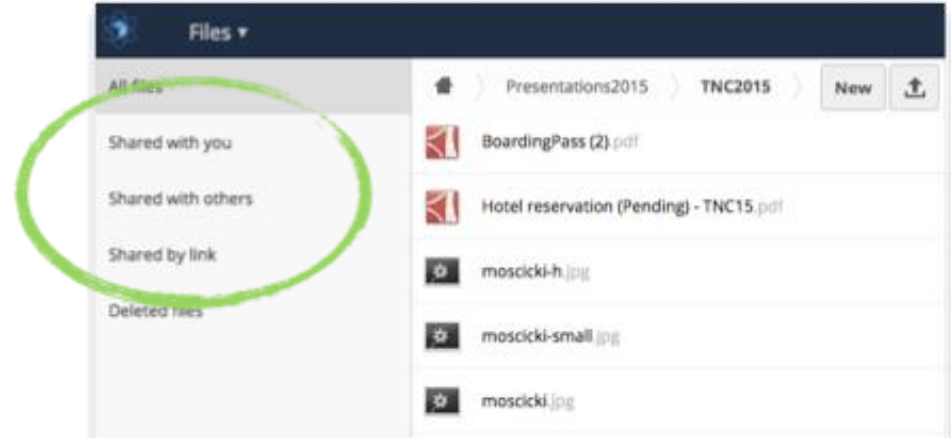


# Access Methods: Sync

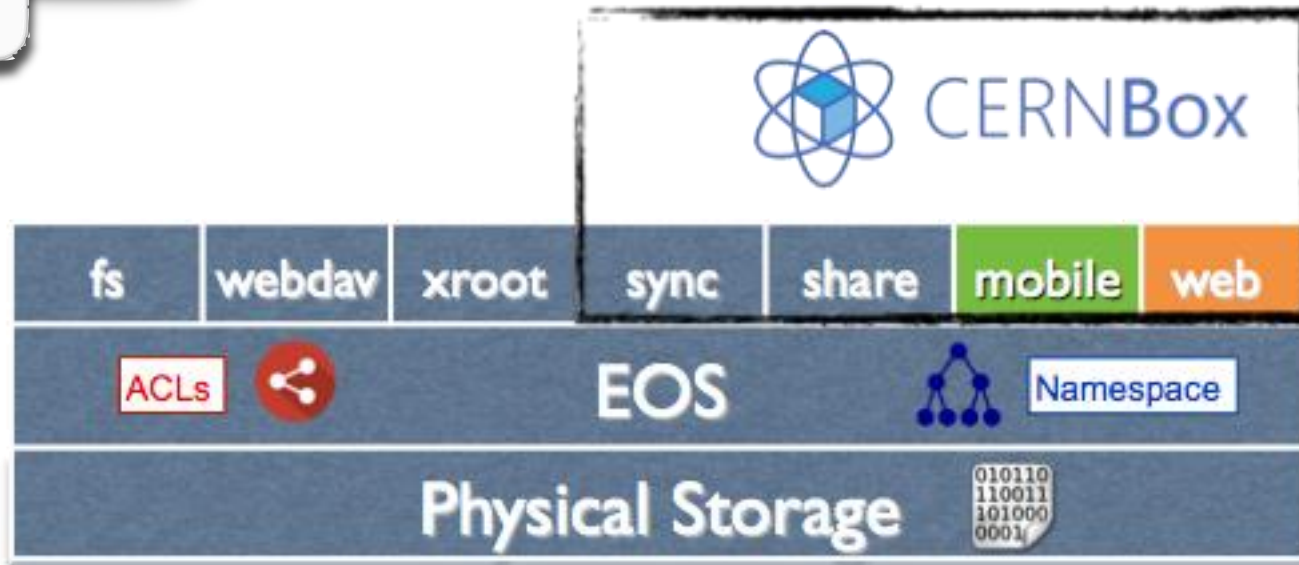
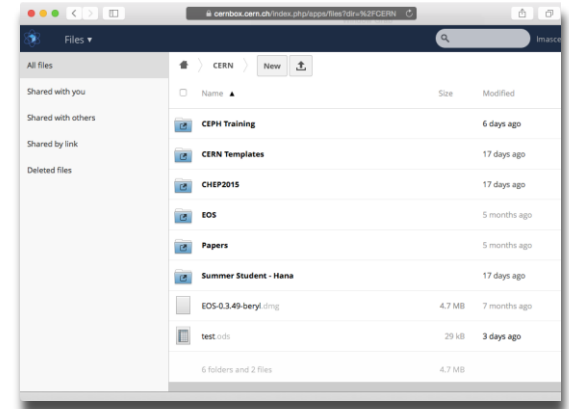




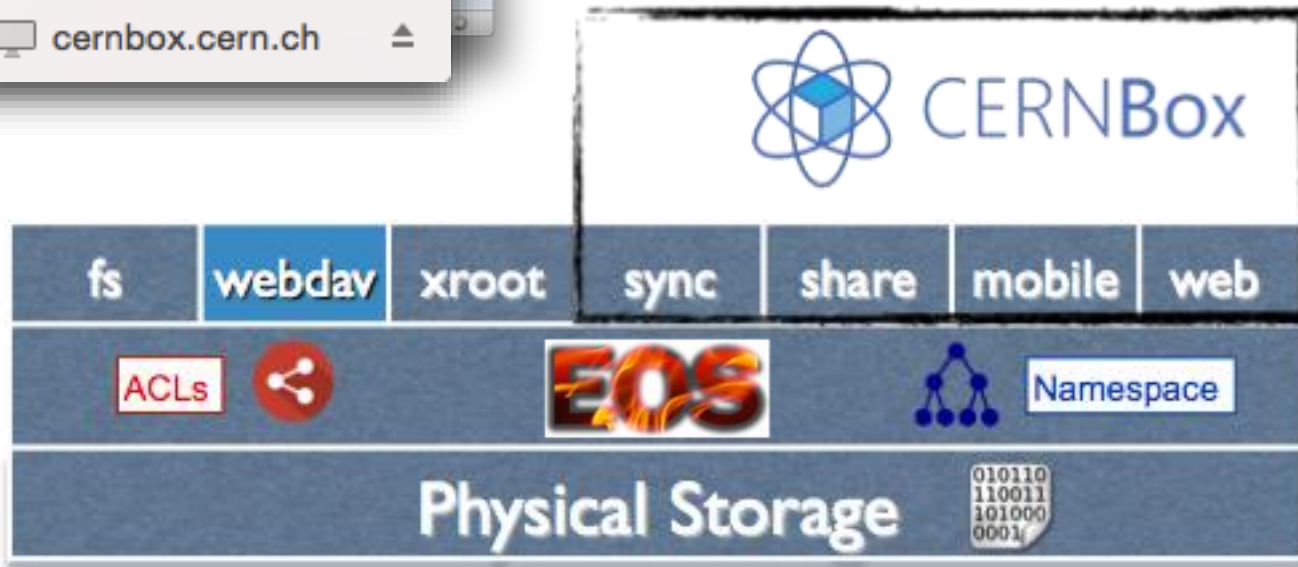
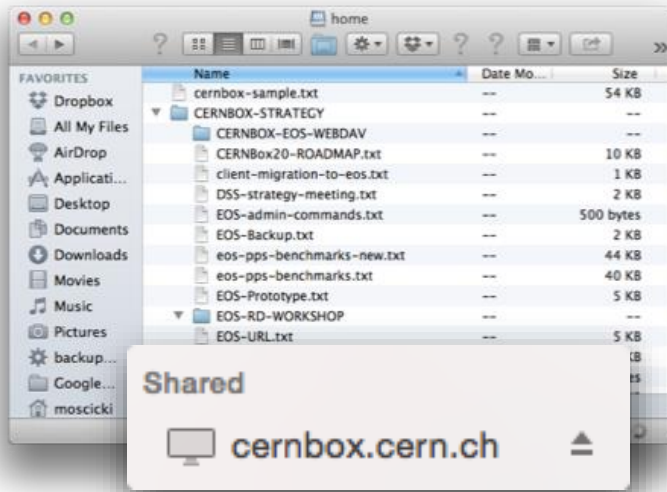
# Access Methods: Sharing



# Access Methods: Mobile & Web



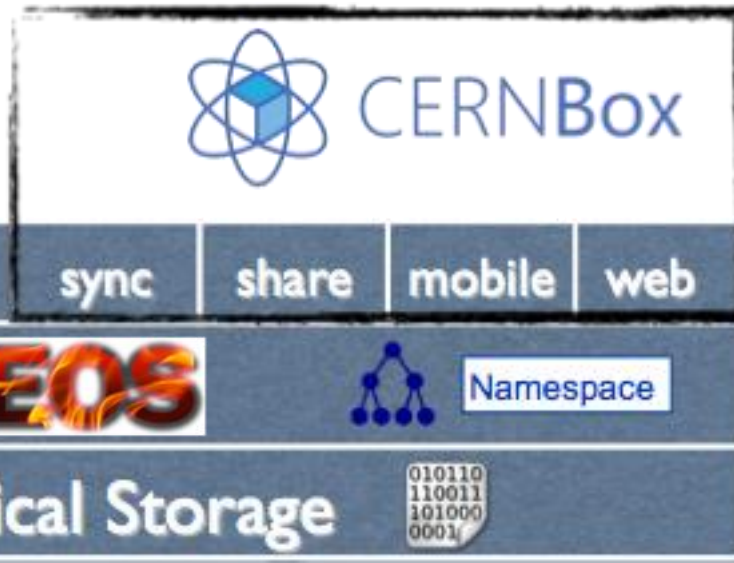
# Access Methods: WebDAV



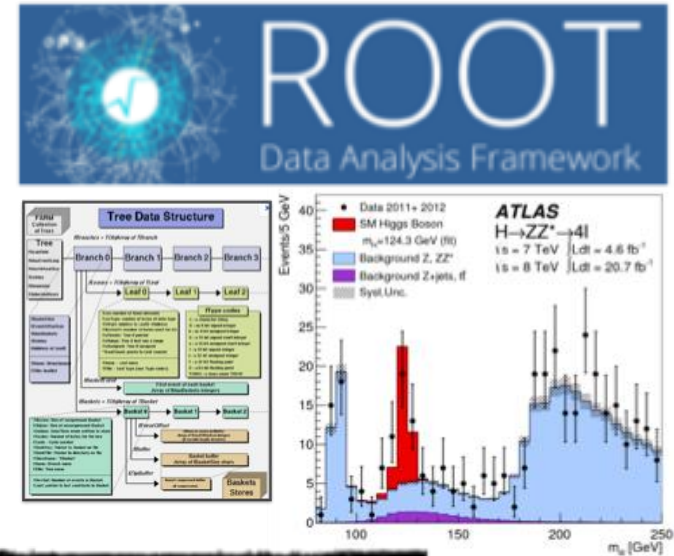
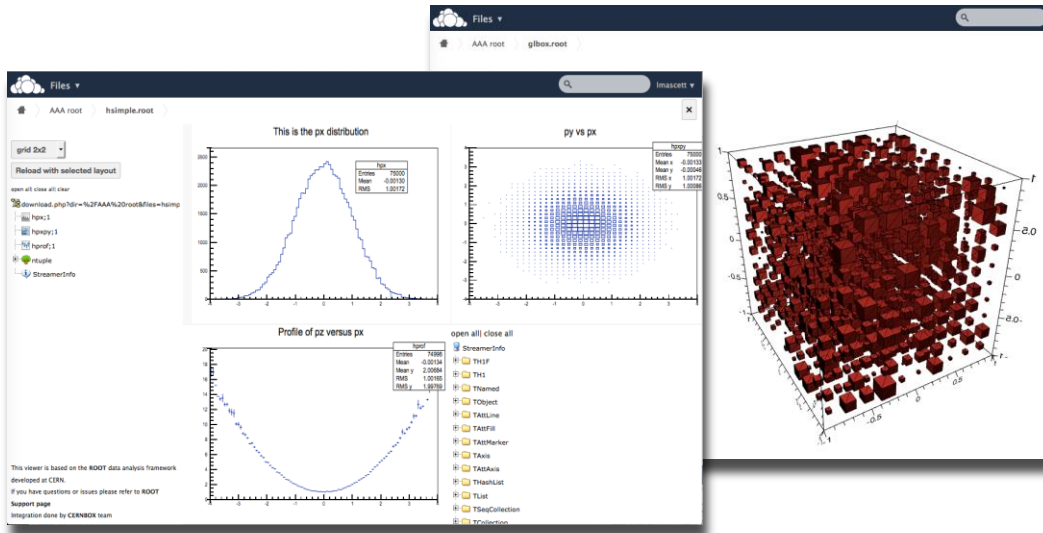
# Access Methods: FUSE



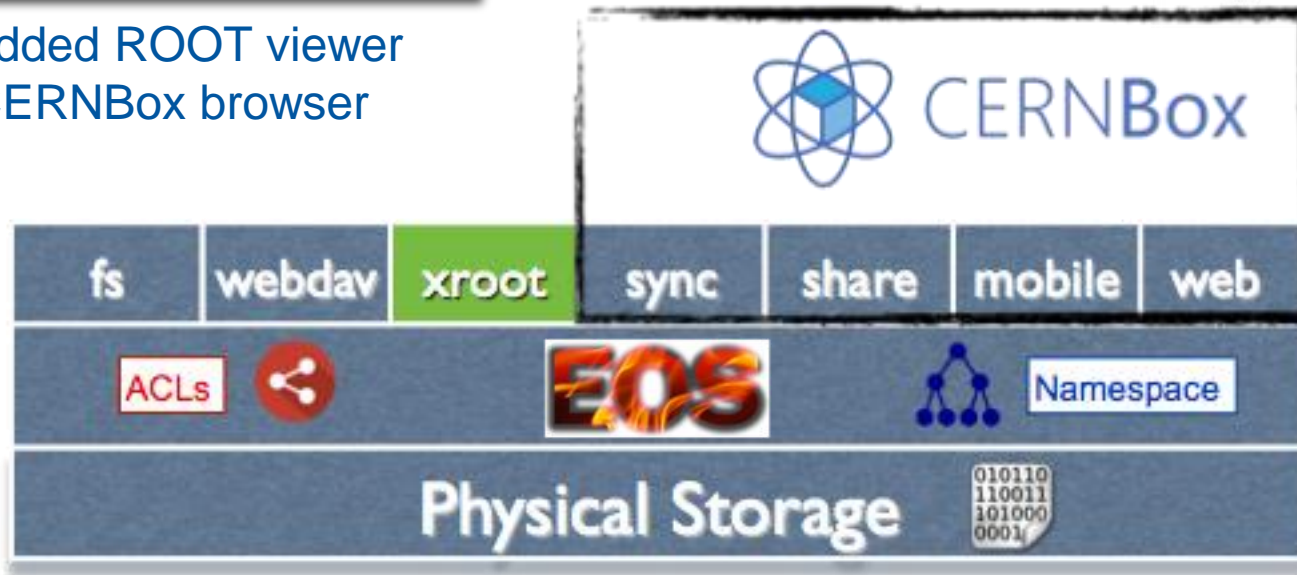
```
[lascett@lxplus2015 ~]#  
[lascett@lxplus2015 ~]# df -H -t fuse  
Filesystem      Size  Used Avail Use% Mounted on  
eosuser         506T  70T  437T  14% /eos/user  
eosatlas        36P   17P   20P  45% /eos/atlas  
eosalice        20P   11P   8.5P  57% /eos/alice  
eoscms          28P   14P   15P  49% /eos/cms  
eoslhcb         13P   7.6P  4.6P  63% /eos/lhcb  
eospublic       16P   5.8P   11P  36% /eos/public  
[lascett@lxplus2015 ~]#  
[lascett@lxplus2015 ~]# ls -lc /eos/user/l/lascett/  
total 6644  
drwx-----. 1 lascett c3      5 Dec 10 15:58 CERN  
drwx-----. 1 lascett c3      0 Jan 26 18:18 debug  
drwx-----. 1 lascett c3      0 Dec 11 09:43 download  
drwx-----. 1 lascett c3      0 Oct 31 18:24 pdf  
drwx-----. 1 lascett c3      1 Dec 11 09:44 personal  
drwx-----. 1 lascett c3      8 Dec 10 12:11 pictures
```




# Optimised access



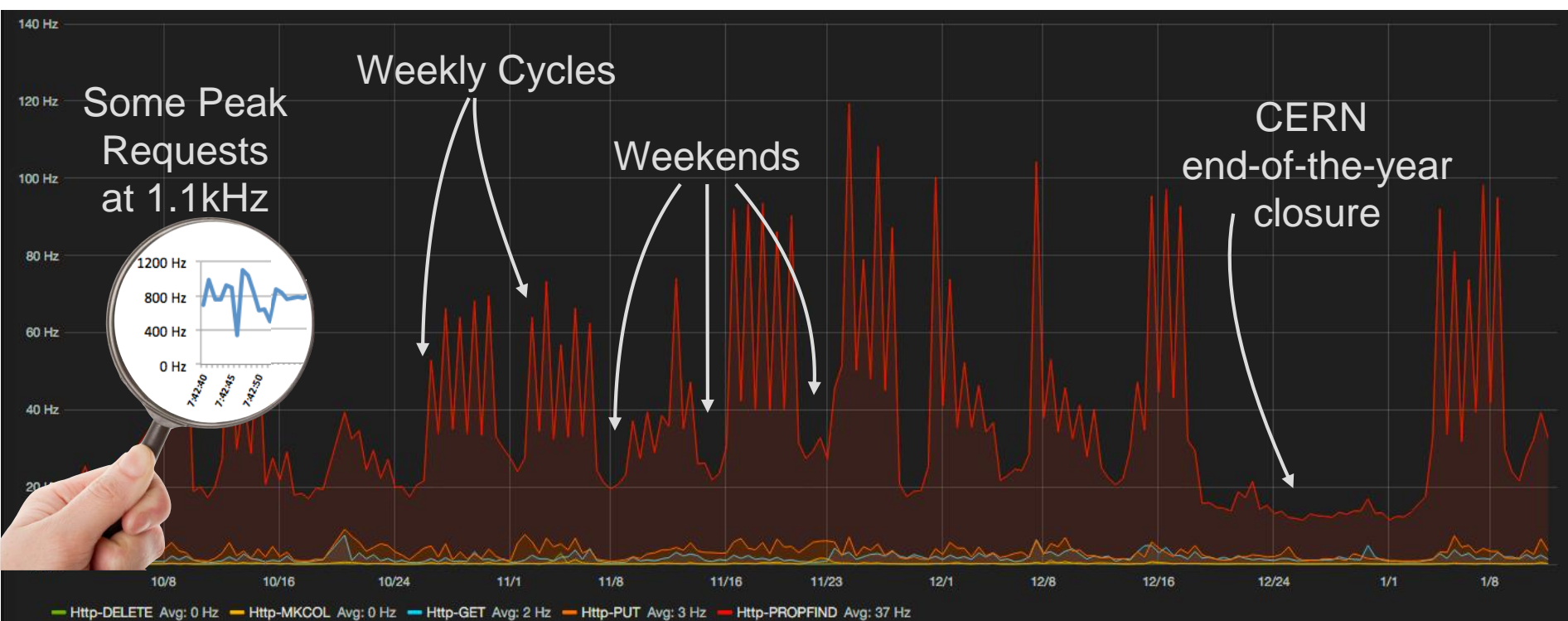
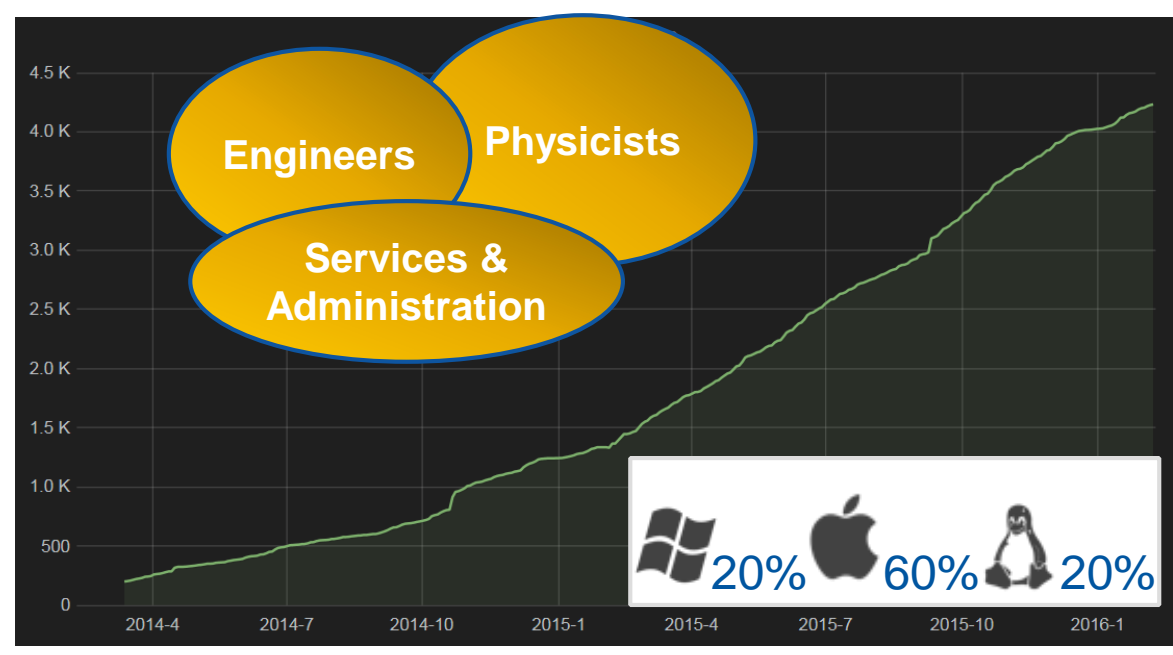
Embedded ROOT viewer  
in CERNBox browser



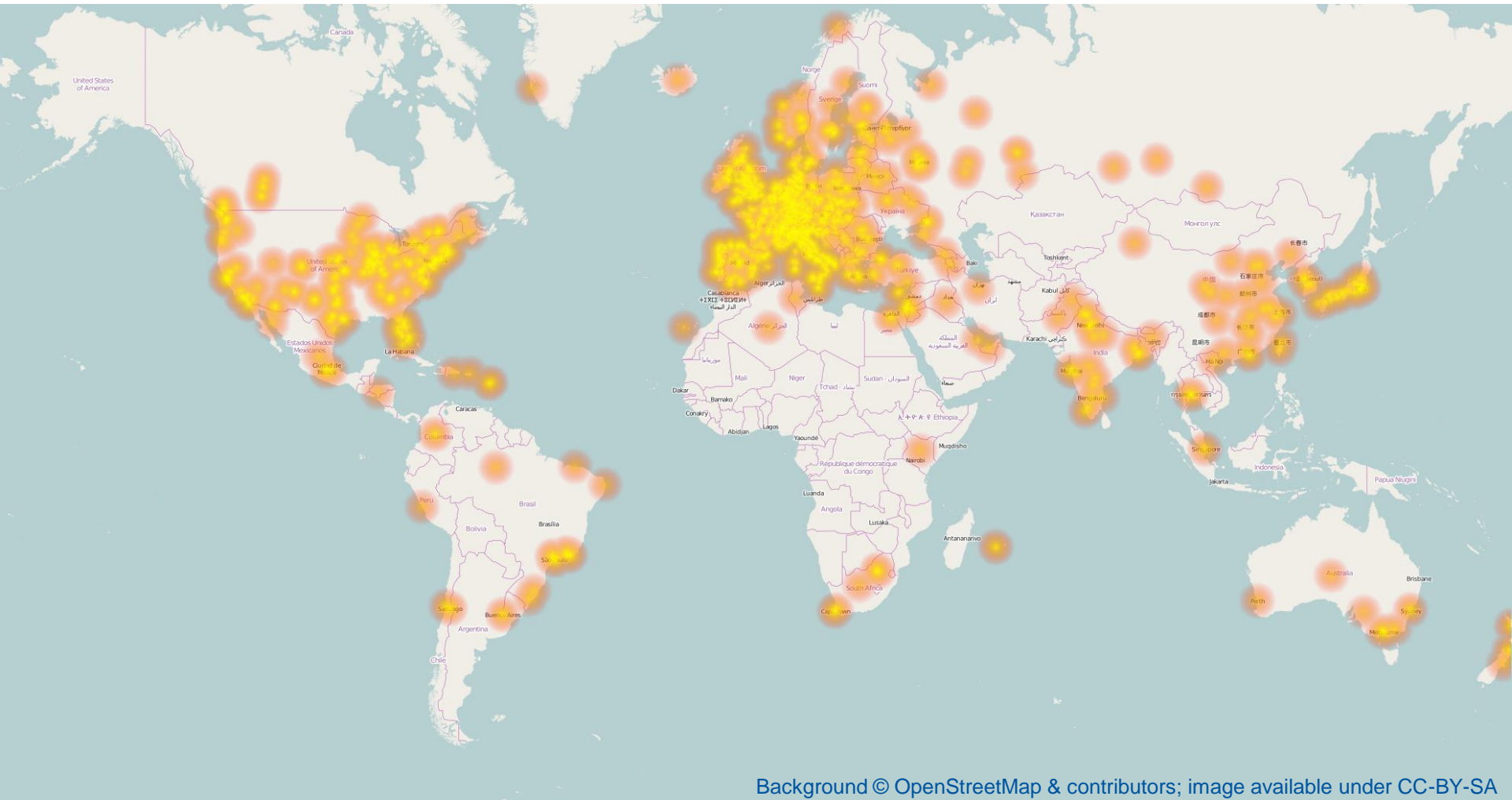
# CERNBox service



<b>Users</b>	<b>4337</b>
<b># files</b>	<b>55 Million</b>
<b># dirs</b>	<b>7.2 Million</b>
<b>Quota</b>	<b>1TB/user</b>
<b>Used Space</b>	<b>104 TB</b>
<b>Deployed Space</b>	<b>1.3 PB</b>



# Dec 2015: Geolocation Active Users

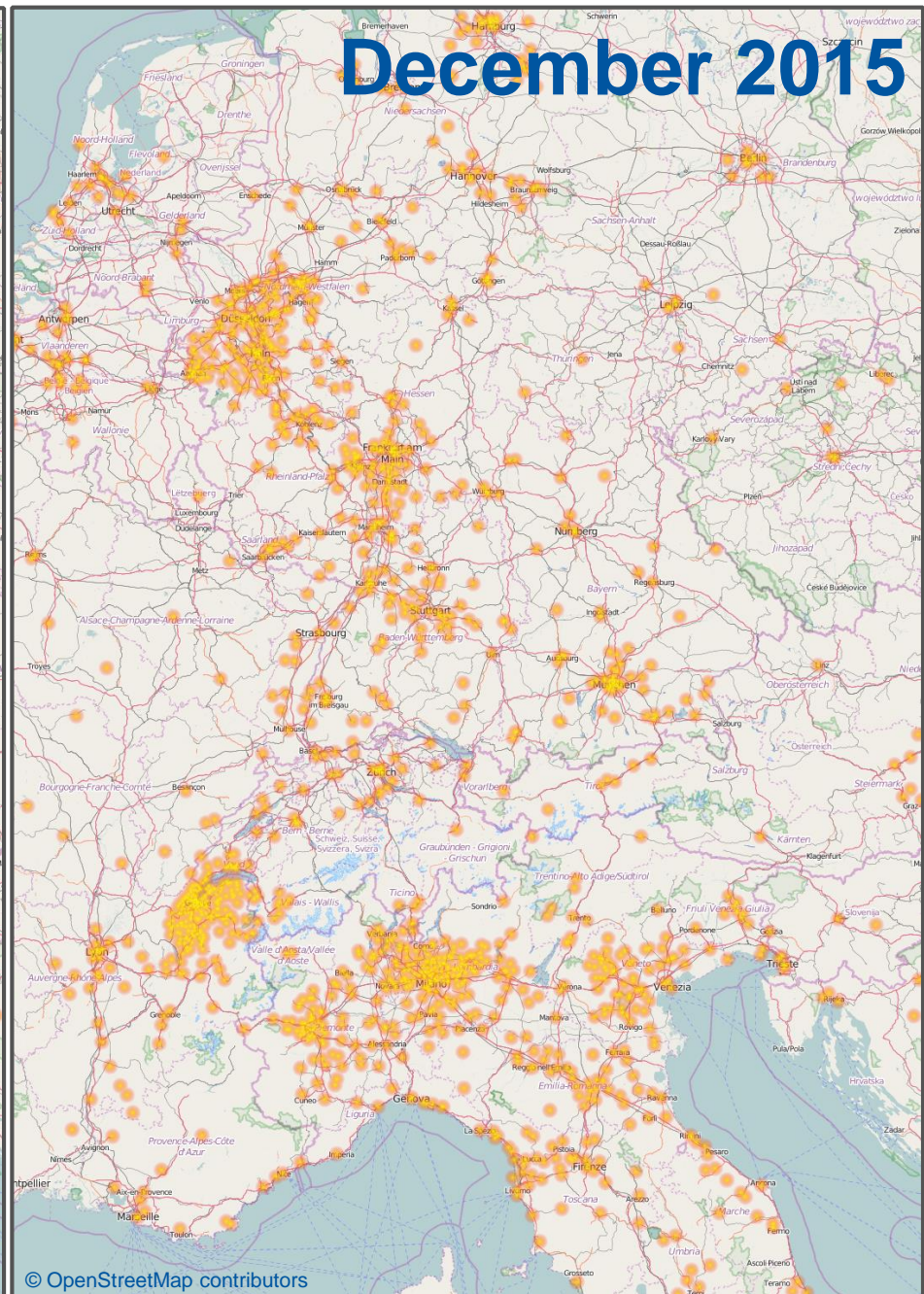
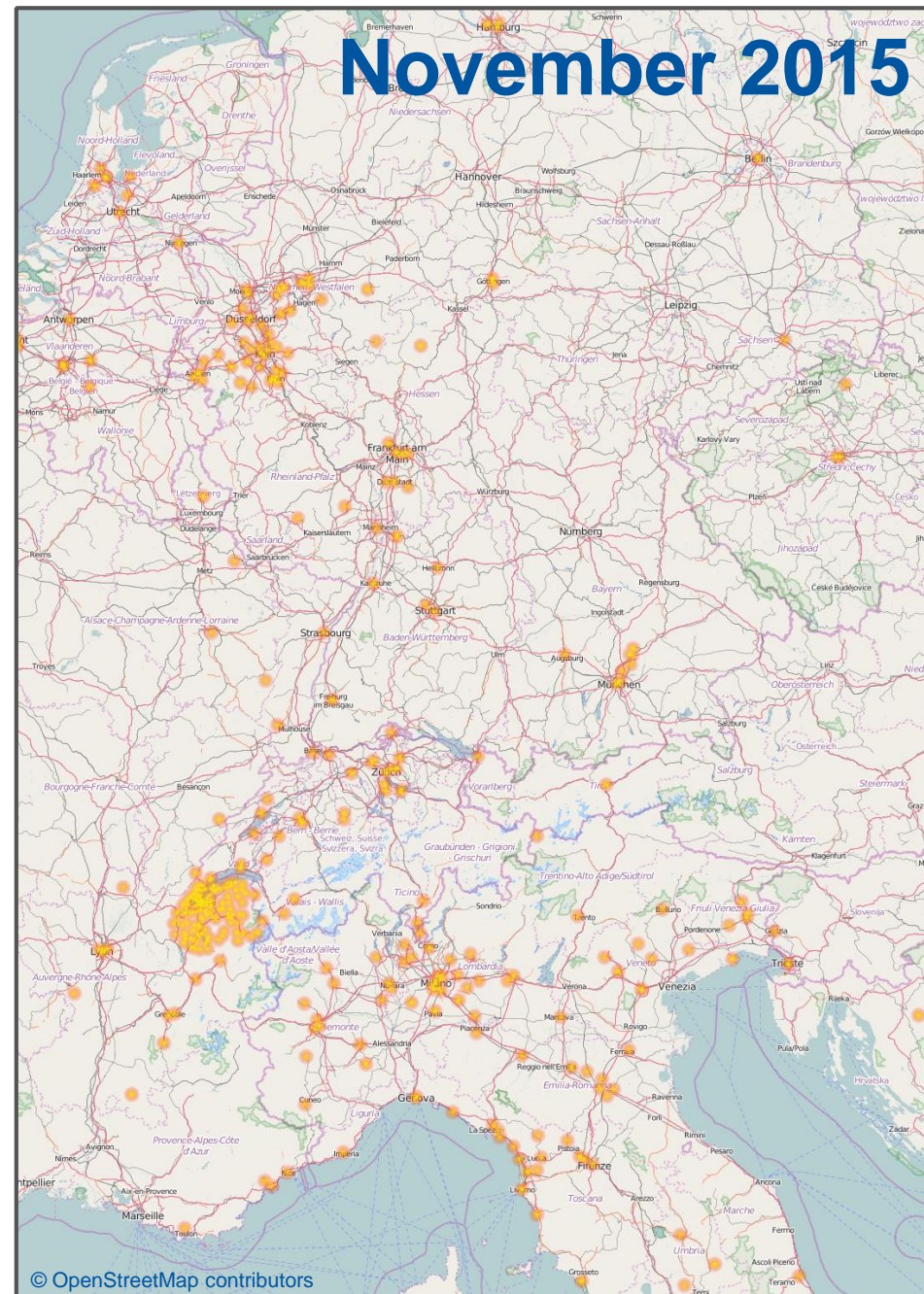


Background © OpenStreetMap & contributors; image available under CC-BY-SA



# November 2015

# December 2015



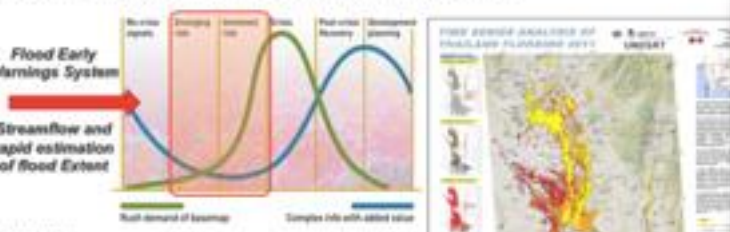
# Usage examples

# CERN Press Office



### A need of a global Flood Early Warning System

Flooding is the most common and widespread hazard worldwide



**Flood Early Warnings System**  
Streamflow and rapid estimation of flood extent

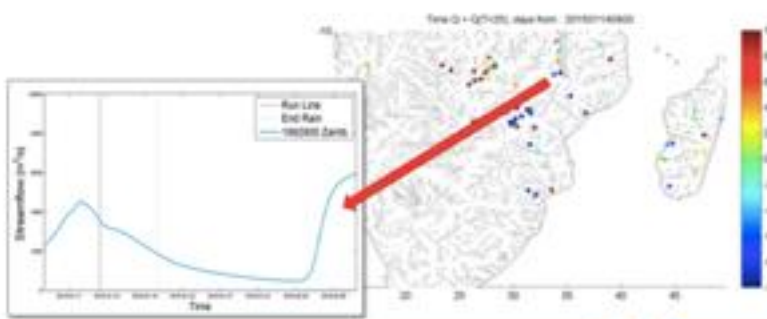
Objectives:

- ✓ Improve disaster response planning with timely identification of potential affected areas, in particular for critical areas of the world that lack of data
- ✓ supporting humanitarian actors during flood emergency with data and analysis
- ✓ guide satellite image acquisition overcome delays due to the triggering process of satellite imagery

### Running hydrological simulations at CERN

With the support of CERN's IT-Department, the operational use of the modelling chain has been tested using CERN grid computing facilities.

1. A service CERN account have been created for the working group,
2. Codes and input data have been uploaded into IT-Dep servers through **CERNBOX**
3. First testing simulation of the forecast modelling chain with the use of IT-Dep Grid Computing platform has been successfully accomplished at CERN on March 2015.



- Enable non-experts to easily use CERN Storage resources
- Powerful integration with the batch system
- Simple to share result with collaborators

# JRC collaboration

## The Joint Research Centre (JRC)

- JRC is the science service of the European Commission
- JRC provides independent scientific support to EU policy making
- Wide usage of **Earth Observation [EO] data** as basis for research and policy support



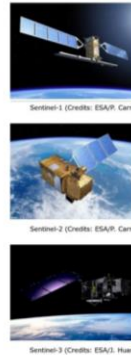
## Proposal for a "JRC Earth Observation Data Processing Platform" (JEO-DPP)

- Main focus on **satellite image** data
- Shall support existing processing workflows and environments (C/C++, Python, Matlab, Java)
- Provide different processing levels:
  - Low-level batch processing
  - High-level interactive processing
- Project timeline:
  - Prototype development: end 2015 – mid 2017
  - Scaling-up in 2017/18: JRC Data Centre vs a public cloud solution



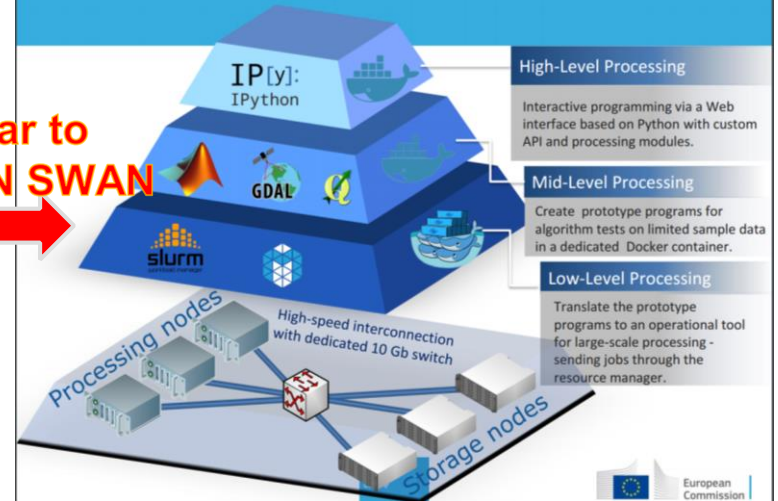
## "Earth Observation & Social Sensing Big Data Pilot Project"

- The EU **Copernicus** Programme with the **Sentinel** fleet of satellites acts as a game changer by bringing EO in the Big Data era:
  - expected 10TB/day of **free and open** data
  - Requires new approaches for data management and processing
- Pilot project launched in January 2015
- Major goal: set up a central infrastructure for storing and processing of Earth Observation and Social Sensing data at JRC



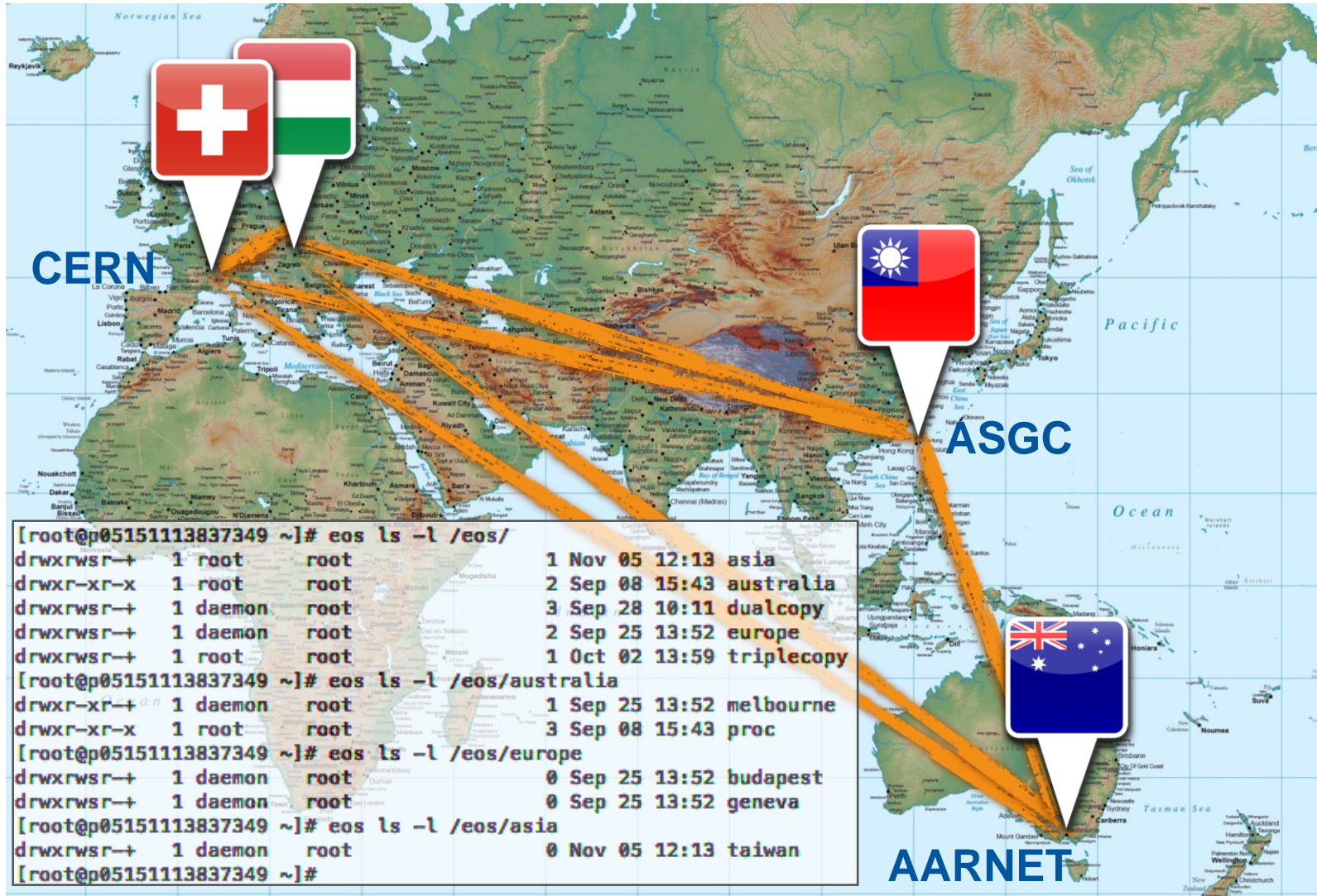
## JEO-DPP processing components

Similar to CERN SWAN

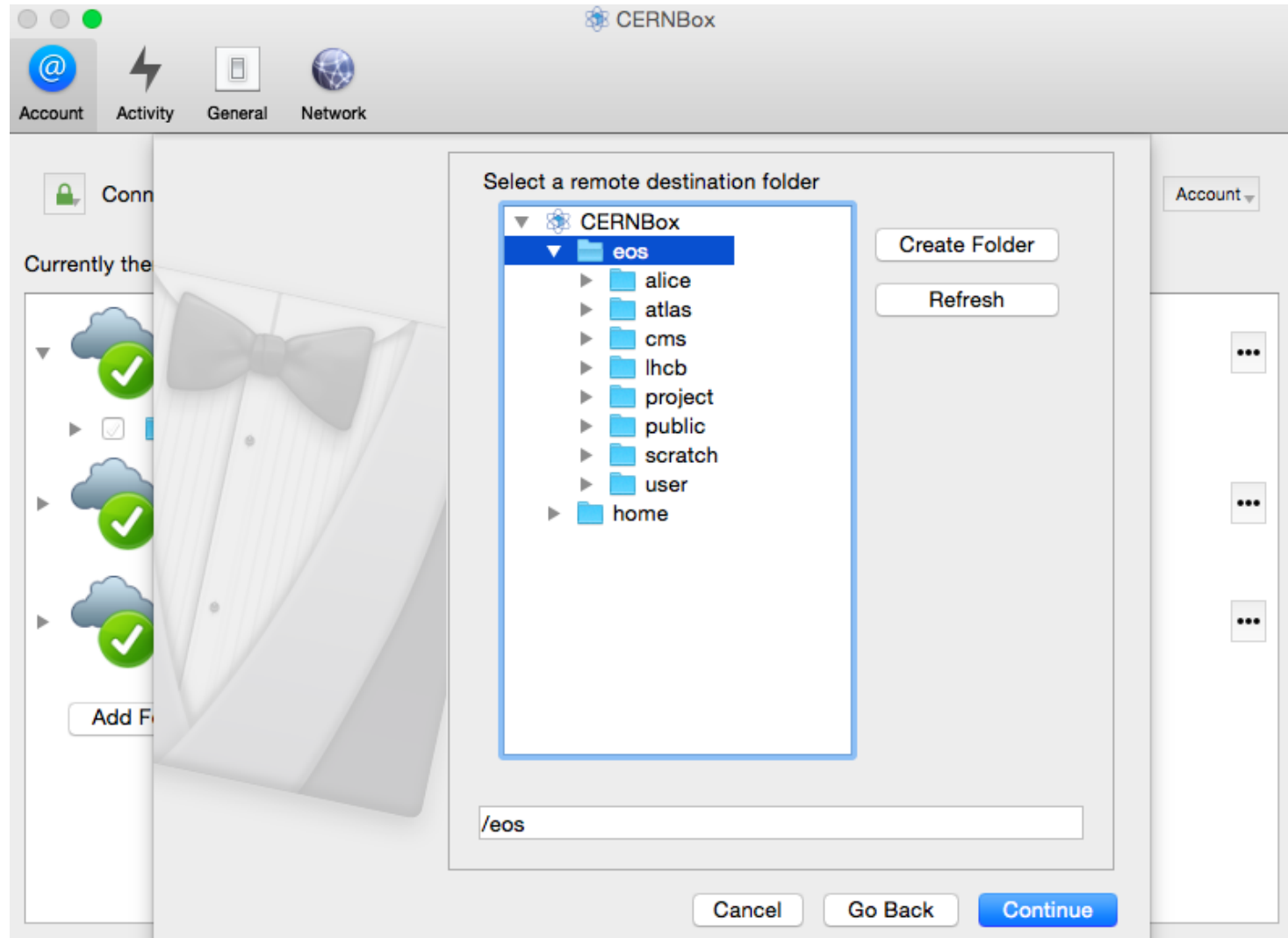


# R&D - EOS World-Wide Deployment

aka "exploring the 300 ms region"...



# CERN data at your fingertips



# Interconnected Private Clouds for Universities and Researchers

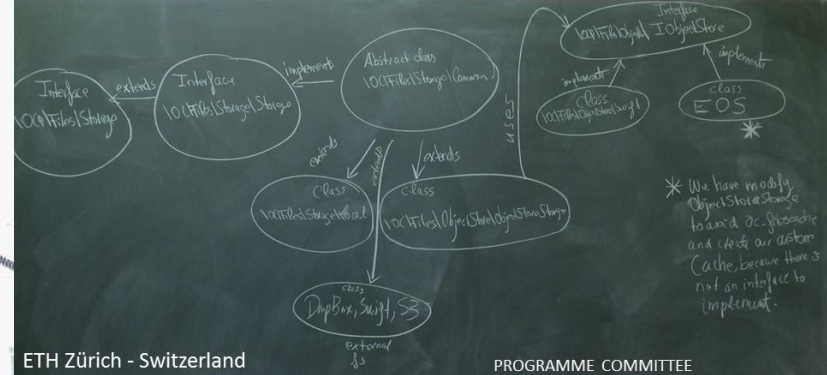


University of Florida  
University of Hawaii - Manoa, College of Education

CERN  
DESY Deutsches Elektronen-Synchrotron  
ERCIS - European Research Center for Information Systems  
ETH Zürich  
Garching Computing Centre of the Max-Planck-Society  
GÉANT Association  
Sciebo, the 500,000 user Campuscloud  
SURF, Netherlands  
SWITCH, Switzerland  
TU Berlin  
University of Vienna

AARNet, Australia

## Cloud Services for Synchronisation and Sharing (CS3) Cloud Storage Services for Novel Applications and Workflows



ETH Zürich - Switzerland  
18-19 January 2016

<https://cs3.ethz.ch>

Abstract submission: 1 September – 15 November 2015

PROGRAMME COMMITTEE  
Massimo Lamanna (CERN), Luca Mascetti (CERN),  
Jakub Mościcki (CERN) and Tilo Steiger (ETH)



# OpenLab project



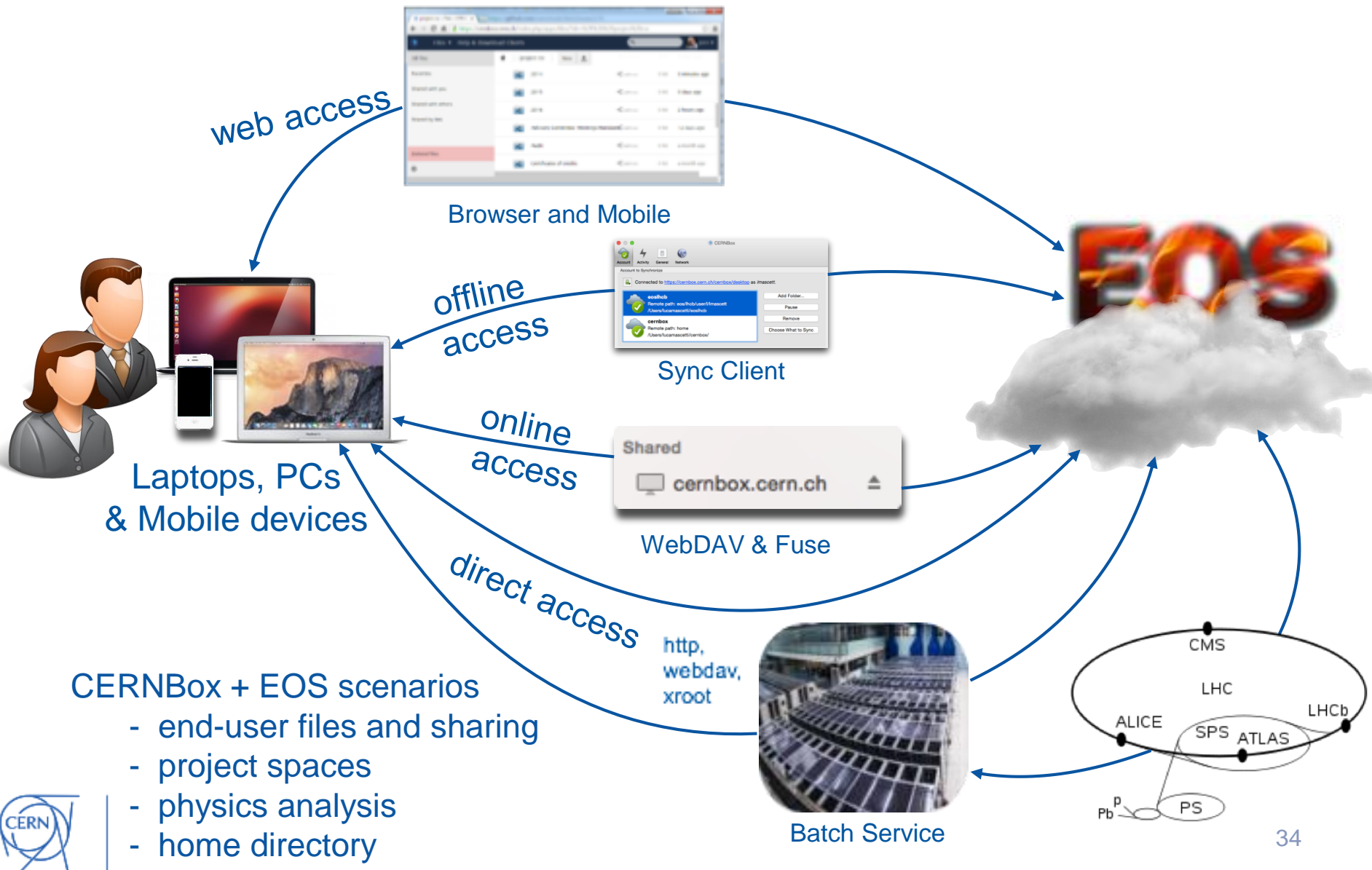
**CERN** openlab

- Streamline EOS
  - More platforms
  - More “generic” installations
    - Remove CERN dependencies
  - More documentation
- Collaboration with **Comtrade**
  - 24-month project (started in July 2015)
  - <http://storage.comtrade.com>
  - Company based in Serbia
  - Technology interest
  - Support opportunity for EOS
    - EOS is open software

The screenshot displays the ComTrade website. At the top left is the ComTrade logo, and to its right is a search bar. Navigation links for 'Expertise', 'Services', 'Technology', 'Solutions', and 'About Us' are visible. The main content area features a headline: 'Internationalization assessment of a leading storage management solution, providing guidelines, effort estimations and remediation schedule plans'. Below this is a navigation bar with several dots, and a sidebar on the left with links for 'Expertise', 'Services', 'Technology', 'Solutions', and 'About Us'. The main text under 'Overview' describes the ComTrade Storage Competence Center and lists key strengths: a thorough understanding of the problem space, wide-ranging and proven knowledge in storage data management, expertise in multi-vendor platforms, and a pool of software engineering talent. A 'Downloads' section on the right lists 'NDMP Server' and 'File System Filler Driver', with a '>NEW' indicator for 'Data Management Suite for Sharepoint'. Contact information for phone and email is also provided at the bottom right.



# Global service for CERN users



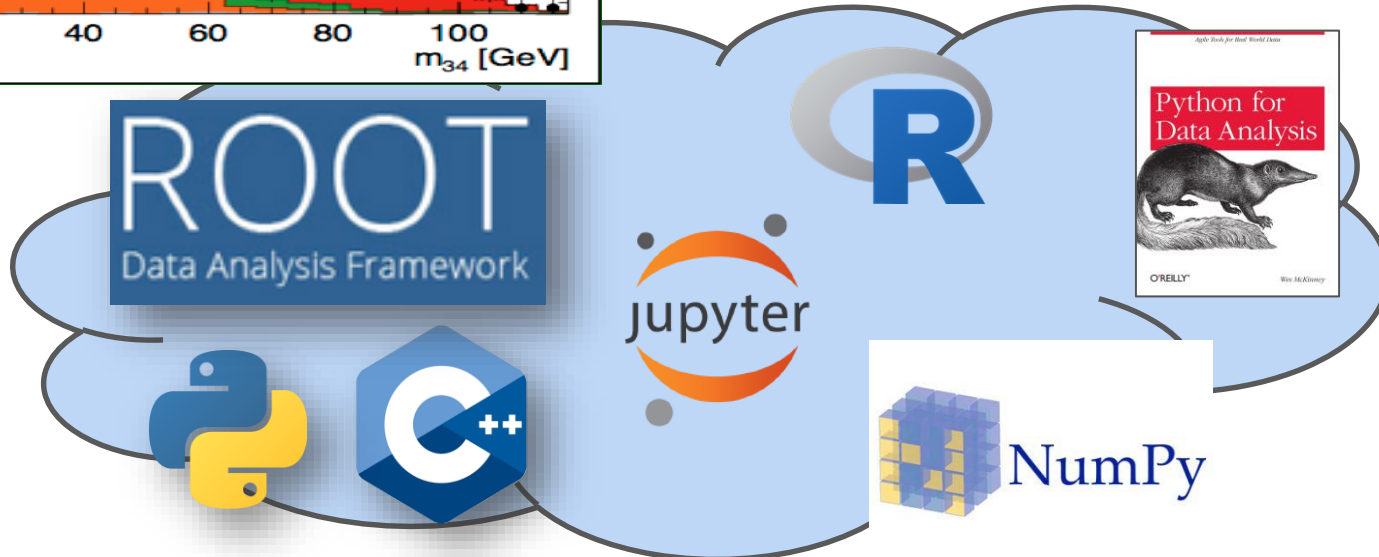
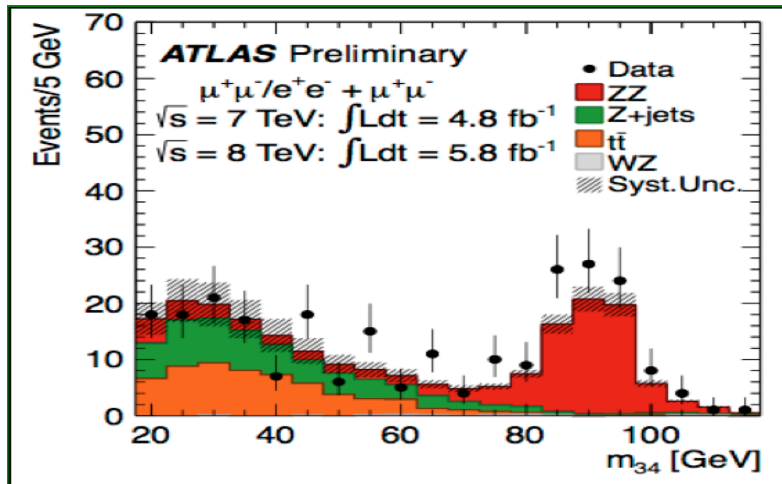
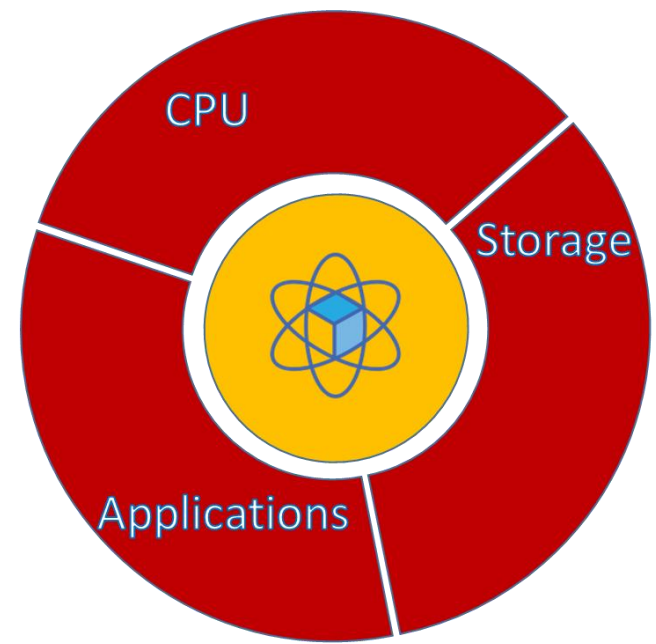
## CERNBox + EOS scenarios

- end-user files and sharing
- project spaces
- physics analysis
- home directory



# Notebook analysis SWAN project

with CERN Physics Department



ROOT is the CERN data analysis framework: <http://root.cern.ch>

Massimo Lamanna | glbox.root - ownCloud | Notebook\_Mano (#1090) | rootbinder/HiggsFit.py

https://cernbox.cern.ch/index.php/apps/files?dir=%2Fdemo%2FR00Tanalysis

Files | Help & Download Clients

demo | ROOTanalysis | glbox.root

simple

Reload with selected layout

open all close all clear

loadfile.php?file=glbox.root&dir=%2Fdemo%2FR00Tanalysis

h31:1

StreamerInfo

## h31

h31	
Entries	5000
Mean x	0.004200
Mean y	0.01368
Mean z	0.002520
Std Dev x	0.5441
Std Dev y	0.5712
Std Dev z	0.5787

This viewer is based on the ROOT data analysis framework, developed at CERN.  
If you have questions or issues please refer to ROOT Support page  
Integration done by CERNBOX team

Search the web and Windows | 17:13 10/04/2016

# Traditional data analysis → Web-based analysis



<https://cernbox.cern.ch/index.php/apps/files?dir=%2Fdemo%2FR00Tanalysis>

Massimo Lamanna | gibox.root - ownCloud | Notebook\_Mano (#10902) | rootbinder/HiggsFit.ipynb

https://cernbox.cern.ch/index.php/apps/files/?dir=%2FNotebook\_Mano%20(%2310902349)

```
import math
PI_Y.append(math.pi) #On defini La valeur de Pi pour L'axe y
PI_Y.append(math.pi) #Et pour L'axe y

c1.cd().SetLogx()

gr = TGraph( n, X, y )
gr.SetLineColor( 1 ) #On choisi La couleur de La ligne
gr.SetLineWidth( 1 ) #Son epaisseur
gr.SetMarkerColor( 4 ) #La couleur des points
gr.SetMarkerStyle( 18 ) #Et Le style de point
gr.SetTitle( 'Methode de Monte Carlo' ) #On choisi Le nom du graphique
gr.GetAxis().SetTitle( 'Nombre de Points' ) #Et Le nom des axes x et y
gr.GetAxis().SetTitle( 'Valeurs estimee de Pi' )

gr.Draw( '' )

gr2=TGraph(2, PI_X, PI_Y)
gr2.SetLineColor( 2 )
gr2.SetLineWidth( 1 )

gr2.Draw("Same")

c1.Draw()
```

TCanvas::Constructor:0: RuntimeWarning: Deleting canvas with same name: c1

Methode de Monte Carlo

On sauvegarde la liste 1 comme ca on peut renouveler l'experience et comparer les resultats.

```
In [25]: list2 = list[:]
```

```
In [26]: list2
```

```
Out[26]: [[10, 3.6, 0.458407346410207],
          [20, 3.0, -0.14159265358979312],
          [30, 3.066666666666667, -0.07492598692312624].
```

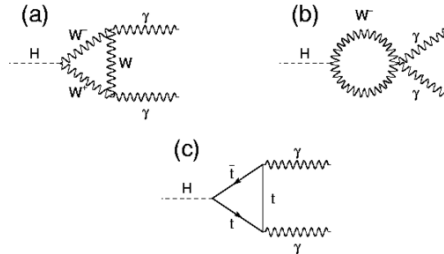
Search the web and Windows | 17:15 10/04/2016

# Training tool to get used with computing, mathematics, ...

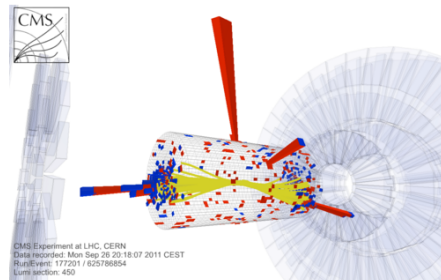


## Higgs decay to two photons

The Standard Model predicted the decay of the [Higgs bosons](#) into photons. The process is depicted by the diagrams below.



At the [Large Hadron Collider](#), this process has been measured. This figure shows how an Higgs boson decay looks in the CMS detector.



This ROOTbook illustrates a simplified fitting procedure aiming to identify the peak due to the Higgs boson decay over the exponentially falling background.

### Importing input data into a ROOT file

First of all we import the input data, here simplistically stored into a text file, into a [ROOT file](#).

```
In [1]: TTree tree("HiggsTree","The tree cont");
auto nevt = tree.ReadFile("Hgg.txt","x");
if (nevt <= 0) {
```

# Outreach to explain our research...



## Higgs decay to two photons

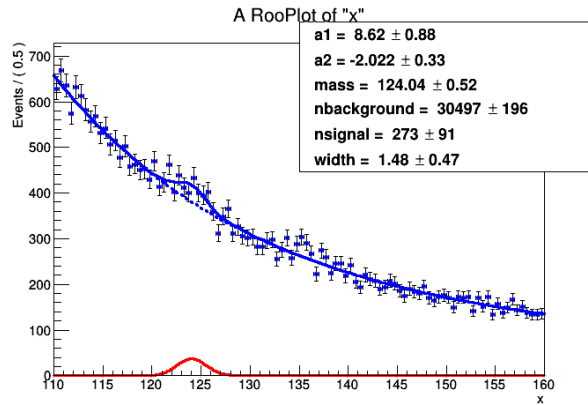
The Standard Model predicted the decay of the [Higgs bosons](#) into photons. The process is depicted by the diagrams below.

(a)  $\gamma$  (b)  $w$

### Visualisation of the Result

```
In [8]: auto plot = x->frame();
data.plotOn(plot);
model->plotOn(plot);
model->plotOn(plot, RooFit::Components("bmodel"), RooFit::LineStyle(kDashed));
model->plotOn(plot, RooFit::Components("smodel"), RooFit::LineColor(kRed));
model->paramOn(plot);

TCanvas c;
plot->Draw();
c.Draw();
```



```
[#1] INFO:NumericIntegration -- RooRealIntegral::init(bmodel_Int[x]) using numeric integrator RooIntegrator
r1D to calculate Int(x)
[#1] INFO:Plotting -- RooAbsPdf::plotOn(model) directly selected PDF components: (bmodel)
[#1] INFO:Plotting -- RooAbsPdf::plotOn(model) indirectly selected PDF components: (z)
[#1] INFO:NumericIntegration -- RooRealIntegral::init(bmodel_Int[x]) using numeric integrator RooIntegrator
r1D to calculate Int(x)
[#1] INFO:Plotting -- RooAbsPdf::plotOn(model) directly selected PDF components: (smodel)
[#1] INFO:Plotting -- RooAbsPdf::plotOn(model) indirectly selected PDF components: ()
[#1] INFO:NumericIntegration -- RooRealIntegral::init(bmodel_Int[x]) using numeric integrator RooIntegrator
r1D to calculate Int(x)
```

### Persistification of the Model on Disk

The ROOT I/O is used to write on disk full RooFit models in order to ease their sharing among scientists, also from different

Full ecosystem with  
CERN tools/initiatives  
like INDICO, Zenodo and  
OpenData

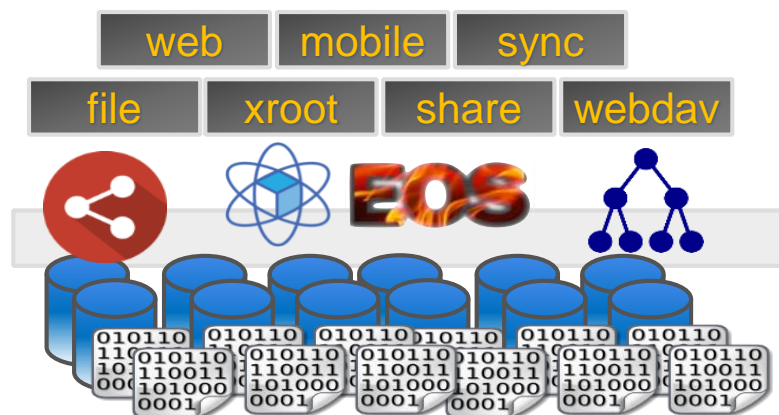


# Outreach to explain our research and get in touch with our data!

# Summary

- Cloud storage enables new use cases
  - and new ways to work and to collaborate
- Solid foundations
  - 200 PB LHC disk infrastructure of EOS (and counting)
  - Data management experience in innovation and service provision
- CERNBox/EOS
  - Home for innovative applications
  - CS3 workshop: visit [cs3.ethz.ch](http://cs3.ethz.ch)

DOI [10.5281/zenodo.44783](https://doi.org/10.5281/zenodo.44783)







[www.cern.ch](http://www.cern.ch)